# The contribution of serial order short-term memory and long-term learning to reading acquisition: A longitudinal study

May 15, 2020

Abstract

There is increasing evidence for an association between both serial order short-term memory (STM) and the long-term learning (LTL) of serial order information and reading abilities. In this developmental study, we examined the hypothesis that STM for serial order supports online grapheme-to-phoneme conversion processes during the initial stages of reading acquisition, whereas the LTL of serial order serves reading abilities at later stages, when reading starts to rely on more stable, long-term orthographic representations. We followed a sample of 116 French-speaking children from first (T1) grade of primary school through second (T2) and third (T3) grade. Their serial order STM and LTL abilities as well as their reading abilities were assessed. Overall, we observed that early reading abilities were only predicted by serial order STM performance, while more advanced reading abilities were predicted by both serial order STM and LTL performance. These results point towards a predictive role of serial order memory performance in reading acquisition and suggest that serial order STM and LTL support reading at different stages of acquisition. We further discuss our findings in the light of advancing knowledge about the relationship between memory and reading.

*Keywords:* Short-term memory; Long-term learning; Serial order; Reading acquisition

Computing spoken language from print in the beginning reader consists of learning the grapheme-phoneme correspondences, which means learning how to pronounce each of the letters and letter combinations of the alphabetic code. When using a nonlexical reading strategy, the reader has to convert each grapheme into its corresponding phoneme. Importantly, to correctly articulate the written word, the sequential product of the grapheme-to-phoneme conversion process may be temporarily maintained and updated until the entire word string has been processed. Repeated decoding of a grapheme sequence by serial reactivation of the corresponding phonemes might then contribute to the gradual creation of a unified orthographic representation in long-term memory for a given written word form. This unified orthographic representation then becomes available for a more direct lexical reading strategy. Learning to read in an alphabetic writing system is therefore assumed to be an order sensitive process (Binamé & Poncelet, 2016c; Bogaerts, Szmalec, De Maeyer, Page, & Duyck, 2016; Martinez Perez, Majerus, & Poncelet, 2012; Whitney, 2001). In other words, the cognitive challenges of reading in alphabetical systems not only lay in learning the regular (and many irregular) grapheme-phoneme correspondences but they also involve learning the exact serial position of the graphemes composing a written word. Alphabetic writing systems are indeed particularly prone to the formation of anagrams (e.g., east, eats, sate, seat, teas), in the sense that thousands of words are all constituted from a limited set of letters. It is therefore important to correctly represent the serial order of the letters composing these anagrams in order to avoid confounding them. More generally, at the beginning of the reading acquisition process, when the child uses a nonlexical letter-by-letter reading strategy, it is essential that all letters composing a novel word are decoded in their correct sequential order. The correct, sequential decoding of a novel word form is a prerequisite for the creation of the corresponding unified lexical representation in long-term memory (Szmalec, Loncke, Page, & Duyck, 2011). Indeed, children who repeatedly inverse the letter order while reading a novel word will require more time to create a stable orthographic representation or may form an erroneous representation. Once a correct orthographic representation of the word has been consolidated in long-term memory, sequential, letter-by-letter decoding, and hence reliance on serial order STM, should diminish. It has been shown indeed that skilled readers, who use a lexical reading strategy, do less readily notice the inversion of two letters in a familiar and frequent word (e.g., “talbe” instead of “table”) compared to beginning readers who still use a nonlexical reading strategy (Grainger, Lete, Bertand, Dufau, & Ziegler, 2012). In line with this reasoning, some authors have proposed that a deficit in the short-term maintenance (Hachmann et al., 2014; Majerus & Cowan, 2016; Martinez Perez, Majerus, Mahot, & Poncelet, 2012) and the long-term learning (LTL) (Bogaerts, Szmalec, Hachmann, Page, & Duyck, 2015; Szmalec et al., 2011) of serial order information can be one of the sources of reading difficulties observed in people with dyslexia, a specific learning deficit that primarily affects the skills involved in accurate and fluent word reading (and/or spelling). In sum, reading can be assumed to involve two important mechanisms: (1) The ability to temporarily represent serial order information during the grapheme-phoneme conversion process (i.e., nonlexical reading), which we suggest to be associated with serial order STM ability and (2) the ability to establish a more durable orthographic representation via long-term serial order learning, which enables fast and efficient reading (i.e., lexical reading strategy).

Two relatively independent lines of research have recently started to investigate the role of verbal serial order memory in reading. One line focused on the involvement of STM for serial order in reading acquisition (Binamé & Poncelet, 2016a, 2016c; Martinez Perez, Majerus, & Poncelet, 2012; Martinez Perez, Poncelet, Salmon, & Majerus, 2015), while the other one explored the link between LTL of sequential information and reading abilities (Bogaerts et al., 2016; Bogaerts et al., 2015; Szmalec et al., 2011).

*The first line of research* relies on recent models of verbal STM that make a distinction between the ability to retain item information (i.e., the identity of the items within a list of items) and the ability to retain serial order information (i.e., the sequential order of the items within the list), thereby typically using immediate serial recall tasks to assess STM performance (Attout, Ordonez Magro, Szmalec, & Majerus, 2018; Burgess & Hitch, 1999, 2006; Gupta, 2003; Majerus & Boukebza, 2013; Majerus & D'Argembeau, 2011; Majerus, Poncelet, Elsen, & Van der Linden, 2006; Majerus, Poncelet, Greffe, & Van der Linden, 2006; Page & Norris, 2009a). The distinction between items and their serial order is critical for understanding the link between STM and language acquisition because item and serial order information are known to interact in different manners with linguistic representations. It has been shown that the recall of item information is directly influenced by existing language knowledge such as phonological, lexical, and semantic representations, while the recall of serial order information is much less influenced by this knowledge. At the same time, serial order learning abilities appear to be a robust predictor of learning novel linguistic representations (Majerus & Boukebza, 2013; Majerus & D'Argembeau, 2011; Nairne & Kelley, 2004; Saint-Aubin & Poirier, 2005). Building on the assumption that recall of serial order reflects a specific memory ability that cannot be reduced to activation of existing language representations (Majerus & Boukebza, 2013; Majerus & D'Argembeau, 2011; Saint-Aubin & Poirier, 2005), the observation of an association between STM for serial order and reading performance thus seems to point towards the existence of a specific and language-independent link between memory capacity and reading acquisition. Some authors have observed that serial order STM is a specific predictor of (later) decoding abilities (Martinez Perez, Majerus, & Poncelet, 2012; Schraeyen, Geudens, Ghesquière, Van der Elst, & Sandra, 2017). More recently, Binamé and Poncelet (2016c) corroborated these results by showing that serial order STM capacity of preschool children is an independent predictor of later nonword reading performance, while it is not related to the ability to read frequent and irregular words. From these observations, the authors deduce that serial order STM may be involved in letter-by-letter decoding processes (nonlexical reading strategies), rather than in direct mapping reading strategies (lexical reading strategies). Further evidence for a link between serial order STM and reading abilities comes from studies on dyslexia (see Majerus & Cowan, 2016, for a review). Martinez Perez, Majerus, Mahot, et al. (2012) as well as Martinez Perez, Majerus, and Poncelet (2013) observed more severe serial order than item STM impairment in children and adults with a history of dyslexia. These observations have been confirmed by a later study in adolescents and adults with dyslexia who showed abnormal activations of a right fronto-parietal network during serial order but not item STM tasks (Martinez Perez et al., 2015). This observation coincides with earlier findings indicating an involvement of the intraparietal sulcus and the dorsolateral prefrontal cortex in serial order STM performance (Majerus et al., 2010).

*The second line of research* reached comparable conclusions concerning the impact of the LTL of serial order information on reading (dis)ability (Bogaerts et al., 2016; Bogaerts et al., 2015; Szmalec et al., 2011). In these studies, serial order LTL was measured via the Hebb repetition paradigm, a commonly used procedure to investigate the LTL of sequential information, which has been shown to engage hippocampal and medial temporal lobe structures (Kalm, Davis, & Norris, 2013). In his seminal work, Hebb (1961) tested participants on an immediate verbal serial recall task in which one particular sequence of digits was repeated every third trial, without the participants knowing. The findings showed that performance for the repeating sequence progressively increases compared to the nonrepeating sequences, which is called the Hebb repetition effect. In other words, this task reflects the transition from a temporary sequence to a more durable representation of this sequence in long-term memory. By using a variant of this Hebb repetition task, Szmalec et al.

(2011) and Bogaerts et al. (2015) observed deficient serial order learning performance in adults with dyslexia. Furthermore, Bogaerts et al. (2016) observed a significant correlation between serial order LTL capacity and nonword reading in a longitudinal study among primary school children. Based on these results, the authors hypothesized that the presence of a serial order learning deficit in reading impaired people could lead to a suboptimal consolidation of the order of the graphemes representing the written words, therefore resulting in impoverished orthographic representations. Although more studies have observed positive correlations between Hebb repetition learning and orthographic abilities in individuals with dyslexia (e.g., Henderson & Warmington, 2016) others have failed to find any such association (Staels & Van den Broeck, 2015, see also Majerus & Cowan, 2016 for a discussion). Despite the considerable amount of studies suggesting that STM and the LTL of serial order information play a role in reading, a yet unexplored question concerns the precise contribution of both serial order STM and LTL to developing reading skills. In other words, it is not clear yet at what point in time serial order STM and LTL are implicated during the reading learning process and how each one precisely supports the cognitive basis of the emerging reading ability. In order to better understand the respective involvement of serial order STM and LTL in reading acquisition, we aimed at bringing both before mentioned lines of research together by contrasting serial order STM with serial order LTL.

In the present study, we used a longitudinal design with 116 primary school children to estimate the predictive role of serial order STM and LTL in reading acquisition. Children were tested in first (T1), second (T2), and again in third grade (T3). At each point in time, we used STM and LTL tasks that were specifically designed to maximize the retention of serial order information. Serial order STM was tested via a serial order reconstruction task designed and validated by Majerus, Poncelet, Greffe, et al. (2006). Serial order LTL on the other hand, was measured through an adaptation of the Hebb repetition learning task (Hebb, 1961). In order to measure reading acquisition as ecologically validly as possible, we decided to measure reading performance with all word types likely to be encountered in French (i.e., regular and irregular words), and we included pseudowords, which can be used to mimic the initial stages of familiarization with an orthographic representation. Reading abilities were assessed at T2 and T3 via the administration of well-established irregular word, regular word, and pseudoword reading tasks. We hypothesized that serial order STM would be particularly implicated in early letter-by-letter decoding processes, allowing for the maintenance of the sequence of phonemes corresponding to the sequence of graphemes of the written word in STM, for subsequent assembly and articulation. On the other hand, we hypothesized that serial order LTL would be specifically involved in the creation of stable orthographic representations in long-term memory, presumably through Hebb repetition learning mechanisms, allowing for a more proficient and automatized reading. In summary, we expected serial order STM to be linked to early reading skills (in second graders), for both words and pseudowords, at a stage where reading abilities mainly draw on letter-by-letter phonological decoding processes. We expected serial order LTL to be linked to reading abilities at T3, when reading starts to be based on unitary orthographic representations consolidated in long-term memory, and this particularly for word reading.

# Method

## Participants

A total of 131 typically developing French-speaking children took part in this longitudinal study. To determine the number of participants, we used the BFDA package (Schönbrodt, 2016) implemented in R (version 3.6.1). The analysis showed that if the correlation of interest exists, the minimal sample size needed for reaching minimal evidence ($BF_{10}>3$) in favor of the correlation in 90% of simulated samples is N=65. For this sensitivity analysis, we assumed an effect size of Pearson's $r$ = .30. This effect size is based on a meta-analysis conducted by Peng et al. (2018), who investigated the association between reading and working memory.

We tested the participants at the end of first grade (T1) ($M_{age}$: 81 months, *SD* = 3.56; 63 girls), at the end of second grade (T2) ($M_{age}$: 93 months, *SD* = 3.59; 63 girls), and again at the end of third grade (T3) ($M_{age}$: 105 months, *SD* = 3.52; 60 girls). Due to a drop-out of 7 children at T2, and of 8 more children at T3, the current analyses were performed on a sample of 116 participants. They were recruited in nine primary schools in the region of Walloon Brabant in Belgium. Parents received a written description of the study as well as a parental and anamnestic questionnaire, which allowed us to ensure that the children were French-speaking monolinguals without a history of neurological disorder, neurodevelopmental delay, sensory, or learning impairment. All children came from families with a middle-class socioeconomic background. All children participated on a voluntary basis and parental consent was obtained. The study has been approved by the ethics committee "Comité d'Éthique Hospitalo-Facultaire de l'Université de Liège" under the project name "Order Memory and Language development" (protocol number 2014-295).

## Materials and procedure

### Tasks administered at 1st and 2nd grade.

#### *Order short-term memory task (Animal Race task).*

This task was an adaptation of the serial order reconstruction task designed and validated by Majerus, Poncelet, Greffe, et al. (2006). It was specifically created to maximize the short-term retention of serial order information, while minimizing item STM requirements. Participants had to listen to sequences of two to seven animal names (chat, chien, coq, lion, loup, ours, singe [cat, dog, cock, lion, wolf, bear, monkey, respectively]) and were subsequently required to rearrange cards with the depicted animals in the order of presentation. As not all animals were used for each trial, at recall the experimenter only provided the cards with the animals involved at presentation. The animal names were announced through headphones at a rate of one item per second. In order to make the task more entertaining for children, we told them that some animals organized a race. The children were asked to arrange the corresponding cards of the animals on a drawn podium (staircase with seven steps on a sheet) in the order they crossed the finishing line. Given the fact that the items of the sequence were available during recall, participants had to reproduce only the serial order of the given sequences and thus, item retention capacity was minimized. The animal names were all monosyllabic, highly familiar and of low age of acquisition, which further allowed to reduce item information processing requirements. For matters of time constraints, we decided to reduce the number of trials per length from four to three trials, relative to the original version of the task (see also Ordonez Magro et al., 2018). In order to avoid a learning effect of the sequences from one year to the other, we created two lists using the same animal names but differing in their sequential order. One list was used at both T1 and T3, while the other one was used at T2. As a dependent variable, we determined the number of correctly placed items by pooling together the different sequence lengths, which gave us a maximum score of 81.

***Hebb repetition learning task.***

In this task, children were instructed to verbally recall sequences of six nonsense consonant-vowel (CV) syllables immediately after auditory presentation. If they forgot one item among the six syllables, they were allowed to say “blanc” (“blank” in English) and to continue the recall. Participants were not made aware of the fact that one sequence, the Hebb sequence, was repeated every second trial, while the other (filler) sequences contained random syllable successions at each trial. The number of six syllables per sequence corresponds to the mean digit span (of 4) in 6-year-olds (Dempster, 1981) plus two more items in order to provide room for progression through repetition learning. Similar to Mosse and Jarrold (2008), we used a non-overlapping design between both types of sequences. This means that the filler sequences consisted of different syllables compared with the Hebb sequences. Both sets of syllables were matched for diphone frequency by using the French database of diphone frequency “Diphones-fr” (New & Spinelli, 2013). The WordGen software was used in order to match all sequences according to their summed diphone frequency (see Table 1; Duyck, Desmet, Verbeke, & Brysbaert, 2004). To avoid sequence-specific effects, we created four different Hebb sequences from one syllable set and counterbalanced them across participants at T1 and T2. The syllable set that served for the creation of the Hebb sequences in first grade became the syllable set for the filler sequences in second grade and vice versa. By doing this, the lists were perfectly matched for diphone frequency. At T3, we created a novel syllable set that was matched for diphone frequency with T1 and T2. For 50% of the participants, half of these syllables were used for the creation of the Hebb sequences, while the other half was used for the creation of the filler sequences, and vice versa for the other 50% of the participants. By doing this, a learning effect across years was avoided. Additionally, we ensured that two or more consecutive syllables never resulted in an existing French word (e.g., SO NA could remind the child of “sauna”). Each participant was presented with 18 sequences: 9 repetitions of the Hebb sequence, interspersed with 9 filler sequences. The syllables were prerecorded by a female voice and presented one at a time through headphones with an inter-stimulus interval of 100 ms. The task always began with a familiarization phase during which the children had to repeat each syllable three times in order to ensure that they could correctly perceive the items. Before starting the learning phase, the children were required to recall one filler sequence, which served as practice trial. Filler and Hebb sequences were presented alternatively as followed: f, H, f, H, f, H, f, H, f … Hebb scores were computed based on a method introduced by McKelvie (McKelvie, 1987; Ordonez Magro et al., 2018; Smalle et al., 2016; Smalle, Page, Duyck, Edwards, & Szmalec, 2017; Staels & Van den Broeck, 2015). This method takes into account both the absolute and relative position of the recalled items. In a first step, the number of items recalled in correct position from left to right up to the first error is determined. In a second step, the same procedure is applied from right to left. In a third step, the number of items recalled in any correct sequence of two or more syllables between the first error from the left and the first error from the right is counted. Finally, any other syllable that occurs in the correct absolute position (independent from its relative position) from left to right is counted. Taking the example "da lu fi pa ve ti ro mi" recalled as "da po fi ve ti mi": "da" would be scored as one item correct (.12) in step 1, "mi" would be scored as one item correct (.12) in step 2, "ve ti" would be scored as two items correct (.25) in step 3, and "fi" would be scored as one item correct (.12) in step 4, as it is recalled in its correct absolute position, even though, relative to its neighbors "lu" and "pa", it is not recalled in the correct position. This would yield a total score of .61.

< INSERT TABLE 1 HERE >

***Receptive vocabulary knowledge.***

In order to measure receptive vocabulary knowledge, we used the EVIP scales (Dunn, Theriault-Whalen, & Dunn, 1993), which is the French adaptation of the Peabody Picture Vocabulary Test (Dunn & Dunn, 1981). Raw vocabulary scores were used for analyses.

***Nonverbal intelligence.***

In order to control for nonverbal intelligence, we used Raven's Progressive Colored Matrices (Raven, Court, & Raven, 1998). Raw scores were used for the analyses.

Estimates of vocabulary knowledge and nonverbal intelligence were collected to ensure that our participant sample was homogeneous with respect to receptive vocabulary or nonverbal intelligence abilities, and that no participant showed developmental delay or cognitive impairment.

**Tasks additionally administered at 2nd and 3rd grade.**

***Reading task.***

This task was designed by Poncelet (1999) and consisted of 30 frequent irregular words, 30 frequent regular words, and 30 pseudowords. Note that the 30 pseudowords used at T2 included contextual graphemes (e.g., in French, the letter “s” is usually pronounced /s/, but when surrounded by vowels, it is pronounced /z/) and thus did not “purely” assess grapheme-to-phoneme mapping abilities but also knowledge of typical French contextual rules represented in procedural memory. In order to test the use of a nonlexical reading procedure in a “purer” manner at T3, we decided to use another list of 30 pseudowords which did not include contextual graphemes and thus were entirely regular at the reading level (e.g., èrlabozi). This list was also created by Poncelet (1999).

The (non)words were presented one by one in a font size of 50 on a 15-inch computer screen using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). The (non)words were printed in lowercase letters and were matched for length (number of letters), lexical frequency and concreteness (for existing words only). Children were asked to read the (non)words as accurately as possible. For each kind of word, we calculated the number of correctly read (non)words.

Note that we also assessed children’s item STM ability, in order to determine the specificity of serial order STM compared to item STM, as predictors of reading performance. Regression analyses showed that item STM only accounts for a negligible part of reading ability as compared to serial order STM, with the consequence that we did not retain the item

STM variable for further analyses. For the interested reader, the methods and full analyses including the item STM variable can be found in the Supplementary Material section.

At each time point, children were tested individually during three sessions in a quiet room at their school. Each session lasted approximately 30 to 40 minutes. At T1, the first session always began with the Hebb task in order to avoid fatigue effects, which was directly followed by the serial order STM task. One day later, the item STM task and the receptive vocabulary task were administered and the order of presentation of both tasks was counterbalanced across participants. During the third session, Raven's colored progressive matrices were administrated in small groups of more or less six children. At T2 and T3, the first session always began with the Hebb task and ended with the serial order STM task. Regarding the second session, we administrated the item STM task, the reading tasks and the receptive vocabulary task, with a counterbalanced order of presentation. Raven's colored progressive matrices were also administrated in small groups of more or less six children during the third session.

We conducted Bayesian analyses, which allows for an unbiased estimation (relative to frequentist statistics) of the effect of interest relative to the null model (Kruschke, Aguinis, & Joo, 2012; Wagenmakers, 2007). We report the $BF_{10}$ as support for the alternative hypothesis ($H_1$) over the null-hypothesis ($H_0$). Small values ($BF_{10}<1$) indicate that there is more evidence for the null hypothesis, and large values ($BF_{10}>1$) indicate more evidence for the alternative hypothesis. Note that we relied on the guidelines proposed by Jeffreys (1961) for interpreting Bayes factors. A BF > 3 provides anecdotal evidence; a BF > 10 provides strong evidence, and a BF > 100 provides decisive evidence for the alternative hypothesis.

## Results

Untransformed descriptive statistics for all tasks are shown in Table 2. T2 regular word and pseudoword reading measures were characterized by a negative skew. Square transformations were applied in order to normalize these measures. Furthermore, as all variables have different maximum scores and ranges, we transformed all scores into z-scores.

< INSERT TABLE 2 HERE >

**Preliminary analyses**

In order to make sure that serial order learning occurred in the Hebb repetition tasks, we first analyzed our Hebb data by collapsing the trials of each sequence type into first- and second half scores (see also Archibald & Joanisse, 2013; Mosse & Jarrold, 2008; Ordonez Magro et al., 2018; Smalle et al., 2016; Smalle et al., 2017). The data on trials 1 to 4 were collapsed into a first half score and the data on trials 6 to 9 were collapsed into a second half score. Serial order LTL was defined in terms of improvement from the first to the second half. Note that Hebb repetition learning reflects the transition from a temporary sequence to a more durable representation of this sequence, and hence reflects both STM and LTL components (Szmalec, Duyck, Vandierendonck, Mata, & Page, 2009). The first score representing performance averaged over the first half of the Hebb trial repetitions can be considered to rely more on STM than on long-term memory processes. The second score representing performance averaged over the second half of the Hebb trial repetitions can be considered to more strongly reflect LTL capacity. We ran a 2 (Sequence type: filler vs. Hebb) x 2 (Half: first vs. second) Bayesian repeated measures ANOVA for T1, T2, and T3. Based on Mosse and Jarrold (2008) as well as Archibald and Joanisse (2013), we compared performance on the first and second halves for each sequence type (Hebb and filler trials). Bayesian analyses were computed by using the JASP software package (Love et al., 2015). Results for the Hebb repetition task at T1, T2, and T3 are plotted in Figure 1.

At T1, results showed decisive evidence for an effect of Sequence type ($BF_{10}$ = $1.58e^{+12}$), with better recall performance for Hebb ($.38 \pm .24_{SD}$) as compared to filler ($.26 \pm .12_{SD}$) sequences. Anecdotal evidence against an effect of Half ($BF_{01} = 1/BF_{10}$, thus,

$BF_{01}$ = 1/0.35 = 2.85) was observed while there was decisive evidence for the presence of an interaction between Sequence type and Half ($BF_{10}$ = 29,985). We observed higher scores on the second half (.43±.27$_{SD}$) than on the first half (.33±.23$_{SD}$) for Hebb but not for filler sequences (.23±.14$_{SD}$ versus .28±.15$_{SD}$), confirming the occurrence of Hebb repetition learning in our sample. Note that the decrease of performance for filler sequences across halves is believed to be due to a fatigue effect (Archibald & Joanisse, 2013; Mosse & Jarrold, 2008).

The same pattern of results was observed at T2. There was decisive evidence for an effect of the Sequence type ($BF_{10}$ = 3.35e$^{+25}$), and substantial evidence against an effect of Half ($BF_{01}$ = 1/0.25 = 4). Finally, there was decisive evidence for the presence of an interaction between Sequence type and Half ($BF_{10}$ = 1,608). The mean proportion of correct responses increased from .44±.24$_{SD}$ for the first half to .54±.28$_{SD}$ for the second half of Hebb repetition trials, while it decreased for filler sequences (.31±.19$_{SD}$ versus .26±.16$_{SD}$).

Regarding T3, we again observe the same pattern of results as in first and second grade. There was decisive evidence for an effect of Sequence type ($BF_{10}$ = 3.05e$^{+33}$), and substantial evidence against an effect of Half ($BF_{01}$ = 7.14). Finally, there was decisive evidence for the presence of an interaction between Sequence type and Half ($BF_{10}$ = 7.66e+8). The mean proportion of correct responses increased from .60±.24$_{SD}$ for the first half to .73±.25$_{SD}$ for the second half of Hebb repetitions, while it decreased for filler sequences (.48±.19$_{SD}$ versus .39±.21$_{SD}$).

In order to compare serial order LTL performance across the three time points, we conducted a Bayesian repeated measures ANOVA with Time (T1 versus T2 versus T3), Sequence type (Hebb versus filler sequences), and Half (first half versus second half) as repeated measures. We obtained decisive evidence for a main effect of Time ($BF_{10}$ = 1.98e$^{+58}$) and Sequence type ($BF_{10}$ = 1.80e$^{+52}$), but anecdotal evidence against a main effect of Half ($BF_{01}$ = 1/0.33 = 3.03). Furthermore, results indicated decisive evidence for an interaction between Time and Sequence type ($BF_{10}$ = 1,167), but strong evidence against an interaction between Time and Half ($BF_{01}$ = 1/0.01 = 59.27). Finally, we observed substantial evidence against an interaction between Time, Sequence type, and Half ($BF_{01}$ = 1/0.14 = 6.79). This last finding indicates that the size of the Hebb effect does not differ across time points.

< INSERT FIGURE 1 HERE >

In sum, we observed a Hebb learning effect at each point of assessment. Note that split-half reliability for filler and Hebb measures across T1, T2, and T3 was reasonable to good (*r* ranging from .54 to .84) and comparable across the three time points (see Table 3). Regarding the serial order STM measure, our participants obtained mean scores similar to those observed in previous studies with children of the same age (Attout et al., 2014; Majerus & Boukebza, 2013). Finally, when correlating serial order STM performance with performance on 2$^{nd}$ half of Hebb sequences after controlling for 2$^{nd}$ half of filler sequences, we observed decisive evidence for an association between serial order STM and serial order LTL at T1 ($r$ = .35, $BF_{10}$ > 100) but only anecdotal evidence for such an association at T2 ($r$ = .19, $BF_{10}$ = 1.22) and T3 ($r$ = .17, $BF_{10}$ = 1.28). The 95% credible interval (comparable to confidence intervals used in frequentist statistics) for the correlation at T1 is [0.18, 0.52] while at T2 it is [0.02, 0.37] and at T3 it is [0.00, 0.35], indicating that the correlation at T1 is robust and clearly different from 0, while this is less the case for the correlations at T2 and T3. These findings point towards a progressive differentiation between serial order STM and LTL mechanisms from T1 to T3.

< INSERT TABLE 3 HERE >

**Bayesian regression analyses**

We conducted Bayesian regression analyses in order to examine the involvement of serial order STM and LTL in reading of irregular, regular, and pseudowords. We anticipated serial order STM to be involved in early reading abilities, when a nonlexical, letter-by-letter sequential reading procedure is mostly used by the beginning reader. Serial order LTL was expected to be associated with later reading abilities, as it is considered to underlie the learning and consolidation of orthographic representations in long-term memory; these representations are established only after repeated reading exposure. Multiple regression analyses assessed the longitudinal predictions between serial order STM/LTL measures and the various reading measures tested one and two years later, while controlling at the same time for contemporaneous serial order memory performance as well as for the influence of nonverbal intelligence (RCPM) and vocabulary knowledge (EVIP)[1]. As a reminder, Hebb repetition tasks reflect both STM and LTL processes (Szmalec et al., 2009). Therefore, we defined serial order LTL as recall performance on the second half of Hebb trial. In the regression analyses, we controlled for serial order STM by introducing the Animal task into the model. We use the term "Hebb measure" to refer to performance on the 2nd half of Hebb trials.

Bayesian regression analysis determines the model (combination of predictor variables) associated with the largest evidence given the data. As a reminder, the bigger the Bayes factor ($BF_{10}$), the stronger the evidence in favor of the corresponding model against the null model (Lee & Wagenmakers, 2013). We present the BF value of the model(s) associated with the strongest evidence and specific Bayesian effect values for each predictor variable ($BF_{forward}$ = comparison of the model with the effect against the null model; $BF_{backward}$ =

1

comparison of the largest model with the effect against the same model that excludes the effect).

For the sake of clarity, the results are presented in the following order: We first give the results regarding the analyses contrasting the implication of T1 STM versus LTL in T2 reading measures. Subsequently, we show the results regarding the analyses contrasting the involvement of T1 STM versus LTL in T3 reading measures. Finally, we present the results regarding the analyses contrasting the implication of T2 STM versus LTL in T3 reading measures.

**T1 serial order STM/LTL and T2 reading.**

***Dependent variable: irregular word reading.***

When contrasting T1 serial order STM and LTL measures, while controlling for contemporaneous T2 serial order memory measures, the predictor variable associated with the highest $BF_{10}$ value against the null model was the T1 serial order STM measure ($BF_{10}$ = 58344), as confirmed by an analysis of specific effects ($BF_{forward}$ = 58344; $BF_{backward}$ = 28). Thus, irregular word reading at T2 was best predicted by serial order STM tested one year earlier. Serial order STM tested at T1 explained 7% of specific variance in T2 irregular word reading.

***Dependent variable: regular word reading.***

When conducting the same set of analyses for T2 regular word reading, the predictor variable associated with the highest $BF_{10}$ against the null model was again the T1 serial order STM measure ($BF_{forward}$ = 3221; $BF_{backward}$ = 7). Similarly to irregular word reading, regular word reading in second graders (T2) was best predicted by T1 serial order STM. T1 serial order STM explained 5% of specific variance in T2 regular word reading scores.

***Dependent variable: pseudoword reading.***

Regarding pseudoword reading, the predictor variable associated with the highest $BF_{10}$ value against the null model was again the T1 serial order STM measure ($BF_{forward}$ = 190; $BF_{backward}$ = 23). As for regular and irregular word reading, pseudoword reading at T2 was best predicted by serial order STM tested at T1. T1 serial order STM explained 8% of the specific variance in T2 pseudoword reading.

In sum, results showed that word and pseudoword reading performances tested at T2 were, as expected, best predicted by T1 serial order STM but not LTL. The findings indicate that early serial order STM is a robust predictor of reading proficiency tested at T2.

**T1 STM/LTL and T3 reading.**

***Dependent variable: irregular word reading.***

When conducting the same set of analyses for T3 irregular word reading measures, while controlling for contemporaneous T3 serial order measures, the predictor variable associated with the highest $BF_{10}$ value against the null model was the T1 serial order STM measure ($BF_{forward}$ = 63283; $BF_{backward}$ = 5). Similarly to word and pseudoword reading at T2, irregular word reading at T3 was best predicted by early serial order STM but not LTL performance. T1 serial order STM explained 7% of the specific variance in T3 irregular word reading.

***Dependent variable: regular word reading.***

Regarding T3 regular word reading, the predictor variable associated with the highest $BF_{10}$ value against the null model was again the T1 serial order STM measure ($BF_{forward}$ = 1596; $BF_{backward}$ = 17). Again, regular word reading at T3 was best predicted by early serial order STM but not LTL ability. T1 serial order STM explained 10% of the specific variance in T3 regular word reading.

***Dependent variable: pseudoword reading.***

About T3 pseudoword reading, the predictor variable associated with the highest $BF_{10}$ value against the null model was once more the T1 serial order STM measure ($BF_{forward}$ = 9166; $BF_{backward}$ = 20). Pseudoword reading at T3 was best predicted by early serial order STM abilities. T1 serial order STM explained 7% of specific variance in T3 pseudoword reading.

In sum, it appears that T3 reading performance is best predicted by early serial order STM abilities, but, contrary to our expectations, not by serial order LTL abilities. These results mirror the already observed findings for the prediction of T2 reading performance.

**T2 STM/LTL and T3 reading.**

***Dependent variable: irregular word reading.***

Considering T3 irregular word reading, the combination of predictor variables associated with the highest $BF_{10}$ value against the null model included the T2 EVIP and the T3 serial order STM measure ($BF_{10}$ = 1.31$e^{+6}$). An analysis of specific effects showed that the variable associated with the largest evidence was the T2 EVIP measure ($BF_{forward}$ = 2958; $BF_{backward}$ = 23) followed by the T3 serial order STM measure ($BF_{forward}$ = 14030; $BF_{backward}$ = 6). The EVIP measure explained 6% and the T3 serial order STM measure explained 5% of variance in T3 irregular word reading. Thus, irregular word reading tested at T3 was best predicted by vocabulary knowledge (EVIP) tested one year earlier and contemporaneous serial order STM ability.

***Dependent variable: regular word reading.***

Regarding T3 regular word reading, the predictor variable associated with the strongest evidence was the T3 LTL measure ($BF_{forward}$ = 61; $BF_{backward}$ = 2). The T3 serial order LTL measure explained 3% of variance in T3 regular word reading.

***Dependent variable: pseudoword reading.***

Finally, regarding T3 pseudoword reading, the predictor variable associated with the highest $BF_{10}$ value against the null model was the T3 serial order STM measure ($BF_{forward}$ = 2490; $BF_{backward}$ = 3). T3 serial order STM explained 4% of variance in pseudoword reading.

The analyses associating T2 serial order memory measures with T3 reading measures revealed that reading performance in third graders was best predicted by *contemporaneous* serial order memory. More precisely, regular word reading was best predicted by contemporaneous serial order LTL, while irregular and pseudoword reading were best predicted by contemporaneous serial order STM. Finally, T3 irregular word reading was also predicted by T2 vocabulary knowledge (i.e., EVIP measure). See Table 4 for a synthesis of results.

< INSERT TABLE 4 HERE >

## Discussion

In the current study, we aimed to directly contrast serial order STM and LTL abilities and to identify their respective associations with reading acquisition. We hypothesized that early reading abilities as measured in second graders (T2) are mostly based on a letter-by-letter sequential decoding strategy which may particularly involve serial order STM mechanisms. Regarding reading abilities in third graders (T3), we expected them to gradually recruit unitary orthographic representations which may be learned via serial order LTL mechanisms. More precisely, we expected serial order STM abilities to longitudinally predict reading of all word types (i.e., irregular, regular, and pseudowords) at T2 and of pseudowords only at T3; we expected serial order LTL abilities to predict familiar word reading but not pseudoword reading abilities at T3. As expected, our findings showed that T2 reading abilities were related to T1 serial order STM but not LTL measures. Regarding T3 reading abilities, regular word reading was linked to contemporaneous (T3) serial order LTL abilities, while irregular word and pseudoword reading was best predicted by contemporaneous serial order STM.

We observed that early T1 serial order STM abilities were linked to all reading measures at both time points, while we expected serial order STM to predict all word types at T2 but only pseudoword reading at T3. A potential explanation why serial order STM was still linked to T3 irregular and regular word reading could be that third graders (T3) still use to some extent a nonlexical reading route to read less familiar words for which orthographic representations are not yet well-represented in long-term memory. As some existing words may still be read by a nonlexical reading route, word reading at T3 may still rely on serial order STM ability. We expect that once reading is fully automatized, the involvement of serial order STM diminishes or disappears completely. The observation that early serial order STM performance as tested at T1 is associated with later reading proficiency is also in line with other studies (Binamé & Poncelet, 2016c; Martinez Perez, Majerus, & Poncelet, 2012). These studies suggested that serial order STM is especially linked to nonlexical reading in young readers because serial order STM supports the temporary maintenance of an ordered phoneme sequence during the grapheme-to-phoneme conversion process until the whole word has been decoded.

In contrast to our hypothesis, early serial order LTL was not related to more advanced reading performances as tested at T3. However, our results showed that a link with T3 regular word reading ability appeared when looking at contemporaneous (T3) serial order LTL ability. Thus, it seems that regular word reading is associated with more advanced serial order LTL abilities. One potential explanation could be that the learning of orthographic representations supporting skilled familiar word reading requires enhanced LTL abilities. This is supported by the stronger Hebb learning performance that is observed at T3. This developmental change in LTL abilities could stem from the use of larger chunks when processing Hebb sequences in older children (Smalle et al., 2016). This developmental increase in chunking abilities may further support skilled reading (like at T3) as it allows for the recognition of progressively larger written segments that define a written word (e.g., rimes, Coltheart & Leahy, 1992). A difference in reliability of the LTL measure at T1, T2, and T3 is not likely to explain our results as this task yielded a similar reliability across the different time points (see Results section). The fact that only regular word reading, but not irregular word reading, was associated to LTL ability could be due to the fact that third graders already rely on established long-term orthographic representations for most regular words, but not yet for all irregular words. This hypothesis is supported by our descriptive statistics showing better reading performance for regular (28/30) compared to irregular (20/30) word reading. A further reason for this result could be that regular words are generally introduced much earlier in the curriculum compared to irregular words (Castles et al., 2018). Thus, our third graders may have encountered regular words more often than irregular words, which may have allowed them to create more stable orthographic representations of regular words, presumably through serial order LTL mechanisms. Furthermore, it should be noted that the irregular words used in this study were composed of rather infrequent orthographic contextual rules. Children at third grade may not yet have sufficient exposure to these orthographic forms to allow learning them via their serial order LTL abilities. Hence, they would still need to rely on a nonlexical reading strategy which is assumed to recruit serial order STM abilities. Finally, the observation that T3 pseudoword reading is linked to contemporaneous serial order STM is in line with our hypothesis that serial order STM specifically supports nonlexical reading processes. The finding that reading measures at T3 are linked to both earlier serial order STM and contemporaneous LTL might indicate that children at third grade use both a nonlexical reading strategy, calling upon serial order STM, as well as a lexical reading strategy, supported by serial order LTL. This assumption is in line with reading models suggesting that nonlexical and lexical reading skills develop simultaneously rather than in stages as soon as children are exposed to written word forms (Grainger et al., 2012; Seidenberg, 2005; Seidenberg & McClelland, 1989; Share, 1995).

Altogether, our findings are in line with the *SERIOL model* proposed by Whitney (2001) that highlights the sequential nature of reading in beginning readers who mostly use a nonlexical reading strategy, which requires the maintenance in STM of the letters constituting the word-to-be-read in their correct serial order while aligning them to their spoken counterparts. Repeatedly processing this letter sequence will then lead to the creation of a unified orthographic representation in long-term memory allowing the use of a more direct lexical reading strategy. Our data showing that serial order STM capacity is linked to early reading performance tested in second graders, while serial order LTL is linked only to more advanced reading tested in third graders (T3), is also in line with recent findings of a study that investigated the role of serial order short- and long-term memory in vocabulary acquisition (Ordonez Magro et al., 2018). This suggests that serial order memory is not only important for reading acquisition but seems to predict in a more general manner the acquisition of lexical knowledge.

A final noteworthy finding of the present study was that irregular word reading at T3 was not only linked to serial order STM and LTL but also to vocabulary knowledge (i.e., EVIP) measured one year earlier. This finding is in line with previous studies showing a link between vocabulary knowledge and irregular word reading (Nation & Snowling, 2004; Peng et al., 2018). According to Ricketts, Nation, and Bishop (2007), children take advantage of their vocabulary knowledge when they fail to correctly decode an irregular word. For example, when encountering the written word “flood”, the child is likely to regularize the word by producing /flud/ and to notice that this is not an existing word. However, by relying on their lexico-semantic representations, they are likely to infer that there is a phonologically similar word to /flud/, which is /flʌd/. This top-down support may thus lead to the correct pronunciation of the irregular word.

As a reminder, the main aim of this study was to directly compare the implication of serial order STM and LTL in *early* (T2) versus *later* (T3) reading ability. However, another interesting question is how serial order STM and LTL abilities can predict *reading growth*. To address this question, we conducted additional regression analyses (see Appendix, Tables A2 and A3) where we represent reading growth by adopting an strategy used in earlier longitudinal studies about the cognitive predictors of reading (e.g., Kyle & Harris, 2010; Lervåg et al., 2009). Like in these studies, we entered earlier levels of reading ability in the regression predicting later reading ability in order to provide "a way to examine the change in reading achievement over time" (Kyle & Harris, p. 236). We also entered contemporaneous serial order STM ability into the model. These analyses indicate that it is especially earlier serial order STM capacity that predicts reading growth, and that pseudoword reading growth is additionally predicted by earlier serial order LTL mechanisms. This last finding is somewhat surprising, as pseudowords are unfamiliar sequences of graphemes that have never been encountered before and that therefore have no representation in long-term memory. Interestingly, similar findings pointing towards a link between serial order LTL and pseudoword reading were already observed by Bogaerts et al. (2016). This observation can be explained by the idea that more experienced readers decode pseudowords by analogy to familiar words, as has been demonstrated before (Coltheart & Leahy, 1992; Ehri & Robbins, 1992). Therefore, we argue that experienced readers are, when decoding novel words, likely to rely on (sub)lexical orthographic representations stored in long-term memory (i.e., segments of familiar words), leading towards an association between serial order LTL and pseudoword reading.

Finally, we would like to note that we only used auditory memory tasks in the present study to measure serial order STM and LTL abilities. Research has shown that the link between serial order memory and (written and spoken) language may exist for both verbal and visuo-spatial serial order memory tasks. Mosse and Jarrold (2008) showed that both verbal and a visuo-spatial Hebb repetition learning tasks predicted novel phonological word learning performance. Furthermore, Martinez Perez et al. (2015) as well as Hachmann et al. (2014) showed that serial order memory deficits, when they occur in individuals with dyslexia, tend to occur in both verbal and visuo-spatial modalities. The assumption of domain-general serial order processing abilities is also supported by neuroimaging findings showing that similar synaptic changes occur regardless of which type of stimuli has to be processed serially (Jensen & Lisman, 2005). In line with the literature, it has been proposed that serial order information is represented in the human memory system in a domain-general manner (Couture & Tremblay, 2006; Depoorter & Vandierendonck, 2009). In line with these studies, we would predict an association between reading acquisition and both auditory and visual serial order STM and LTL tasks, a hypothesis to be tested in future studies.

## Conclusion

The present longitudinal study aimed to investigate the nature of the association between serial order STM/LTL mechanisms and reading acquisition. Our findings indicate that early reading abilities, based on letter-by-letter phonological decoding processes are linked to earlier serial order STM abilities while more advanced reading performances are supported by contemporaneous LTL abilities, with additional support from serial order STM abilities. This study suggests that serial order STM and LTL support reading development via distinct but complementary processes.

## References

Archibald, L. M. D., & Joanisse, M. F. (2013). Domain-specific and domain-general constraints on word and sequence learning. *Memory & Cognition, 41*(2), 268-280. doi:10.3758/s13421-012-0259-4

Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science*, *36*(2), 286-304. https://doi.org/10.1111/j.1551-6709.2011.01200.x

Attout, L., & Majerus, S. (2015). Working memory deficits in developmental dyscalculia: The importance of serial order. *Child Neuropsychol, 21*(4), 432-450. doi:10.1080/09297049.2014.922170

Attout, L., Noël, M. P., & Majerus, S. (2014). The relationship between working memory for serial order and numerical development: A longitudinal study. *Developmental Psychology, 50*(6), 1667-1679. doi:http://dx.doi.org/10.1037/a0036496

Attout, L., Ordonez Magro, L., Szmalec, A., & Majerus, S. (2018). The developmental neural substrates of item and serial order components of verbal working memory. *Hum Brain Mapp*. doi:10.1002/hbm.24466

Biname, F., & Poncelet, M. (2016a). The development of the abilities to acquire novel detailed orthographic representations and maintain them in long-term memory. *Journal of Experimental Child Psychology, 143*, 14-33. doi:http://dx.doi.org/10.1016/j.jecp.2015.10.010

Biname, F., & Poncelet, M. (2016b). The development of the abilities to acquire novel detailed orthographic representations and maintain them in long-term memory. *J Exp Child Psychol*, *143*, 14-33. doi:10.1016/j.jecp.2015.10.010

Biname, F., & Poncelet, M. (2016c). Order short-term memory capacity predicts nonword reading and spelling in first and second grade. *Reading and Writing, 29*(1), 1-20. doi:10.1007/s11145-015-9577-9

Bogaerts, L., Szmalec, A., De Maeyer, M., Page, M. P., & Duyck, W. (2016). The involvement of long-term serial-order memory in reading development: A longitudinal study. *Journal of Experimental Child Psychology, 145*, 139-156. doi:10.1016/j.jecp.2015.12.008

Bogaerts, L., Szmalec, A., Hachmann, W. M., Page, M. P., & Duyck, W. (2015). Linking memory and language: Evidence for a serial-order learning impairment in dyslexia. *Research in Developmental Disabilities*, *43-44*, 106-122. doi:10.1016/j.ridd.2015.06.012

Burgess, N., & Hitch, G. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106*(3), 551-581. doi:10.1037/0033-295X.106.3.551

Burgess, N., & Hitch, G. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language, 55*(4), 627-652. doi:10.1016/j.jml.2006.08.005

Castles, A., Rastle, K., & Nation, K. (2018). Ending the Reading Wars: Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest*, *19*(1), 5-51. https://doi.org/10.1177/1529100618772271

Coltheart, V., & Leahy, J. (1992). Children's and adults' reading of nonwords: Effects of regularity and consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(4), 718-729. doi:http://dx.doi.org/10.1037/0278-7393.18.4.718

Couture, M., & Tremblay, S. (2006). Exploring the characteristics of the visuospatial Hebb repetition effect. *Memory & Cognition, 34*(8), 1720-1729. doi:http://dx.doi.org/10.3758/BF03195933

Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin, 89*(1), 63-100. doi:10.1037/0033-2909.89.1.63

Depoorter, A., & Vandierendonck, A. (2009). Evidence for Modality-Independent Order Coding in Working Memory. *Quarterly Journal of Experimental Psychology, 62*(3), 531-549. doi:10.1080/17470210801995002

Dunn, L. M., & Dunn, L. M. (1981). *Peabody picture vocabulary test–revised: Manual for forms L and M.* Circle Pines, MN: American Guidance Service.

Dunn, L. M., Theriault-Whalen, C. M., & Dunn, L. M. (1993). *Échelle de vocabulaire en images Peabody:Adaptation française du Peabody Picture Vocabulary Test.* Toronto, Canada: Psycan.

Duyck, W., Desmet, T., Verbeke, L. P. C., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments, & Computers, 36*(3), 488-499. doi:10.3758/bf03195595

Ehri, L. C., & Robbins, C. (1992). Beginners Need Some Decoding Skill to Read Words by Analogy. *Reading Research Quarterly, 27*(1), 13-26. JSTOR. https://doi.org/10.2307/747831

Grainger, J., Lete, B., Bertand, D., Dufau, S., & Ziegler, J. C. (2012). Evidence for multiple routes in learning to read. *Cognition, 123*(2), 280-292. doi:10.1016/j.cognition.2012.01.003

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology*

*A: Human Experimental Psychology, 56A*(7), 1213-1236. doi:10.1080/02724980343000071

Hachmann, W. M., Bogaerts, L., Szmalec, A., Woumans, E., Duyck, W., & Job, R. (2014). Short-term memory for order but not for item information is impaired in developmental dyslexia. *Ann Dyslexia*, *64*(2), 121-136. doi:10.1007/s11881-013-0089-5

Hebb, D. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37-46). Oxford, UK: Blackwell.

Henderson, L. M., & Warmington, M. (2016). A sequence learning impairment in dyslexia? It depends on the task. *Research in Developmental Disabilities*. doi:http://dx.doi.org/10.1016/j.ridd.2016.11.002

Jeffreys, H. (1961). *Theory of probability* (3 ed.). Oxford: Oxford University Press, Clarendon Press.

Jensen, O., & Lisman, J. E. (2005). Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. *Trends Neurosci, 28*(2), 67-72. doi:10.1016/j.tins.2004.12.001

Kalm, K., Davis, M. H., & Norris, D. (2013). Individual sequence representations in the medial temporal lobe. *Journal of Cognitive Neuroscience, 25*(7), 1111-1121. doi:10.1162/jocn_a_00378

Kay, J., & Marcel, A. (1981). One Process, Not Two, in Reading Aloud: Lexical Analogies Do the Work of Non-Lexical Rules. *The Quarterly Journal of Experimental Psychology Section A, 33*(4), 397-413. doi:10.1080/14640748108400800

Kruschke, J. K., Aguinis, H., & Joo, H. (2012). The Time Has Come:Bayesian Methods for Data Analysis in the Organizational Sciences. *Organizational Research Methods*, *15*(4), 722-752. doi:10.1177/1094428112457829

Kyle, F. E., & Harris, M. (2010). Predictors of reading development in deaf children : A 3-year longitudinal study. *Journal of Experimental Child Psychology*, *107*(3), 229-243. https://doi.org/10.1016/j.jecp.2010.04.011

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*: Cambridge University Press, New York, NY.

Lervåg, A., Bråten, I., & Hulme, C. (2009). The cognitive and linguistic foundations of early reading development : A Norwegian latent variable longitudinal study. *Developmental Psychology*, *45*(3), 764-781. https://doi.org/10.1037/a0014132

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., & Verhagen, A. J. e. a. (2015). JASP (Version 0.7).

Majerus, S., & Boukebza, C. (2013). Short-term memory for serial order supports vocabulary development: new evidence from a novel word learning paradigm. *Journal of Experimental Child Psychology, 116*(4), 811-828. doi:10.1016/j.jecp.2013.07.014

Majerus, S., & Cowan, N. (2016). The nature of verbal short-term impairment in dyslexia: The importance of serial order. *Frontiers in Psychology, 7,* 8.

Majerus, S., & D'Argembeau, A. (2011). Verbal short-term memory reflects the organization of long-term memory: Further evidence from short-term memory for emotional words. *Journal of Memory and Language, 64*(2), 181-197. doi:10.1016/j.jml.2010.10.003

Majerus, S., D'Argembeau, A., Perez, T. M., Belayachi, S., Van der Linden, M., Collette, F., . . . Maquet, P. (2010). The commonality of neural networks for verbal and visual short-term memory. *Journal of Cognitive Neuroscience, 22*(11), 2570-2593. doi:http://dx.doi.org/10.1162/jocn.2009.21378

Majerus, S., Poncelet, M., Elsen, B., & Van der Linden, M. (2006). Exploring the relationship between new word learning and short-term memory for serial order recall, item recall, and item recognition. *European Journal of Cognitive Psychology, 18*(6), 848-873. doi:10.1080/09541440500446476

Majerus, S., Poncelet, M., Greffe, C., & Van der Linden, M. (2006). Relations between vocabulary development and verbal short-term memory: The relative importance of short-term memory for serial order and item information. *Journal of Experimental Child Psychology, 93*(2), 95-119. doi:10.1016/j.jecp.2005.07.005

Martinez Perez, T., Majerus, S., Mahot, A., & Poncelet, M. (2012). Evidence for a specific impairment of serial order short-term memory in dyslexic children. *Dyslexia: An International Journal of Research and Practice, 18*(2), 94-109. doi:http://dx.doi.org/10.1002/dys.1438

Martinez Perez, T., Majerus, S., & Poncelet, M. (2012). The contribution of short-term memory for serial order to early reading acquisition: Evidence from a longitudinal study. *Journal of Experimental Child Psychology, 111*(4), 708-723. doi:http://dx.doi.org/10.1016/j.jecp.2011.11.007

Martinez Perez, T., Majerus, S., & Poncelet, M. (2013). Impaired short-term memory for order in adults with dyslexia. *Research in Developmental Disabilities, 34*(7), 2211-2223. doi:http://dx.doi.org/10.1016/j.ridd.2013.04.005

Martinez Perez, T., Poncelet, M., Salmon, E., & Majerus, S. (2015). Functional alterations in order short-term memory networks in adults with dyslexia. *Developmental Neuropsychology, 40*(7-8), 407-429. doi:http://dx.doi.org/10.1080/87565641.2016.1153098

McKelvie, S. J. (1987). Learning and awareness in the Hebb digits task. *Journal of General Psychology, 114*(1), 75-88. doi:10.1080/00221309.1987.9711057

Mosse, E. K., & Jarrold, C. (2008). Hebb learning, verbal short-term memory, and the acquisition of phonological forms in children. *The Quarterly Journal of Experimental Psychology, 61*(4), 505-514. doi:10.1080/17470210701680779

Nairne, J. S., & Kelley, M. R. (2004). Separating item and order information through process dissociation. *Journal of Memory and Language, 50*(2), 113-133. doi:10.1016/j.jml.2003.09.005

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading, 27*(4), 342-356. doi:http://dx.doi.org/10.1111/j.1467-9817.2004.00238.x

New, B., & Spinelli, E. (2013). Diphones-fr: A French database of diphone positional frequency. *Behavior Research Methods, 45*(3), 758-764. doi:http://dx.doi.org/10.3758/s13428-012-0285-y

Ordonez Magro, L., Attout, L., Majerus, S., & Szmalec, A. (2018). Short-and long-term memory determinants of novel word form learning. *Cognitive Development, 47*, 146-157. doi:https://doi.org/10.1016/j.cogdev.2018.06.002

Page, M., & Norris, D. (2009a). A model linking immediate serial recall, the Hebb repetition effect and the learning of phonological word forms. *Philos Trans R Soc Lond B Biol Sci, 364*(1536), 3737-3753. doi:10.1098/rstb.2009.0173

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., . . . Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin, 144*(1), 48-76. doi:http://dx.doi.org/10.1037/bul0000124

Poncelet, M. (1999). Exploration du rôle des composants phonologique et visuel de la mémoire à court terme dans l’apprentissage des procédures de lecture (Unpublished doctoral thesis). University of Liege.

Raven, J. C., Court, J. H., & Raven, J. (1998). *Progressive matrices couleur*. Oxford, UK: Oxford Psychologists Press.

Ricketts, J., Nation, K., & Bishop, D. V. M. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading, 11*(3), 235-257. doi:http://dx.doi.org/10.1080/10888430701344306

Sawi, O. M., & Rueckl, J. (2018). Reading and the neurocognitive bases of statistical learning. *Scientific Studies of Reading*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1080/10888438.2018.1457681

Saint-Aubin, J., & Poirier, M. (2005). Word frequency effects in immediate serial recall: Item familiarity and item co-occurrence have the same effect. *Memory, 13*(3-4), 325-332. doi:10.1080/09658210344000369

Schönbrodt, F. D. (2016). BFDA: An R package for Bayes Factor Design Analysis, version 0.1. Https://github.com/nicebread/BFDA.

Schraeyen, K., Geudens, A., Ghesquière, P., Van der Elst, W., & Sandra, D. (2017). Poor performance on the retention of phonemes' serial order in short-term memory reflects young children's poor reading skills. *The Mental Lexicon, 12*(1), 129-158. doi:https://doi.org/10.1075/ml.12.1.06sch

Seidenberg, M. S. (2005). Connectionist models of word reading. *Current Directions in Psychological Science, 14*(5), 238-242. doi:http://dx.doi.org/10.1111/j.0963-7214.2005.00372.x

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychol Rev, 96*(4), 523-568.

Share, D. L. (1995). Phonological Recoding and Self-Teaching: Sine qua non of Reading Acquisition. *Cognition, 55*(2), 151-218.

Smalle, E. H. M., Bogaerts, L., Simonis, M., Duyck, W., Page, M. P., Edwards, M. G., & Szmalec, A. (2016). Can Chunk Size Differences Explain Developmental Changes in Lexical Learning? *Frontiers in Psychology, 6*, 1925. doi:10.3389/fpsyg.2015.01925

Smalle, E. H. M., Page, M. P., Duyck, W., Edwards, M., & Szmalec, A. (2017). Children retain implicitly learned phonological sequences better than adults : A longitudinal study. *Developmental Science*.

Staels, E., & Van den Broeck, W. (2015). No solid empirical evidence for the solid (serial order learning impairment) hypothesis of dyslexia. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(3), 650-669. doi:10.1037/xlm0000054

Szmalec, A., Duyck, W., Vandierendonck, A., Mata, A. B., & Page, M. P. A. (2009). The Hebb repetition effect as a laboratory analogue of novel word learning. *The Quarterly Journal of Experimental Psychology, 62*(3), 435-443. doi:10.1080/17470210802386375

Szmalec, A., Loncke, M., Page, M. P. A., & Duyck, W. (2011). Order or disorder? Impaired Hebb learning in dyslexia. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(5), 1270-1279. doi:http://dx.doi.org/10.1037/a0023820

Tubach, J. L., & Boe, L. J. (1990). *Un corpus de transcription phonétique.* France: Telecom.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review, 14*(5), 779-804. doi:http://dx.doi.org/10.3758/BF03194105

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review, 8*(2), 221-243. doi:http://dx.doi.org/10.3758/BF03196158

## Author Note

Laura Ordonez Magro, Laboratory of Cognitive Psychology, Aix-Marseille University; Psychological Sciences Research Institute, Université catholique de Louvain

Steve Majerus, Psychology and Neuroscience of Cognition Research Unit, Université de Liège

Lucie Attout, Psychology and Neuroscience of Cognition Research Unit, Université de Liège

Martine Poncelet, Psychology and Neuroscience of Cognition Research Unit, Université de Liège

Eleonore H.M. Smalle, Department of Experimental Psychology, Ghent University

Arnaud Szmalec, Psychological Sciences Research Institute and Institute of Neuroscience, Université catholique de Louvain; Department of Experimental Psychology, Ghent University

This work was supported by a grant from the Fonds de la Recherche Scientifique – FRS – FNRS (PDR, FRFC, T.1003.15, Belgium).

Correspondence concerning this article should be addressed to Laura Ordonez Magro, Laboratory of Cognitive Psychology, Aix-Marseille University.

E-mail: laura.ordonezmagro@gmail.com

## Acknowledgements

The authors thank Agathe Leroy, Marie Gomrée, Pauline Scalliet, and Tamara Neve de Mévergnies for their help in the data collection and all the school directors, teachers, pupils, and parents for their time and collaboration.

## Footnote

[1] We entered contemporaneous serial order memory performance as a control variable in our regression analyses in order to control for the possibility that the relation between serial order memory and later reading performance is just a consequence of a correlation between early and later serial order memory ability.

Table 1

*Stimulus material for the Hebb tasks at T1, T2, and T3. French diphone frequency for each syllable is reported.*

| Filler (T1)/ Hebb (T2) | | Hebb (T1)/ Filler (T2) | | Filler/ Hebb (T3) | | Filler/ Hebb (T3) | |
|---|---|---|---|---|---|---|---|
| CV | Diphone | CV | Diphone | CV | Diphone | CV | Diphone |
| TI [ti] | 3440 | RI [ʁi] | 3880 | LI [li] | 2843 | SI [si] | 2199 |
| PA [pa] | 1755 | MI [mi] | 1670 | DI [di] | 1825 | MA [ma] | 1814 |
| FI [fi] | 1142 | NA [na] | 1262 | TA [ta] | 2137 | RO [ʁɔ] | 1947 |
| VE [və] | 50 | BE [bə] | 29 | PE [pə] | 72 | NE [nə] | 117 |
| DA [da] | 497 | GU [gy] | 173 | VO [vɔ] | 308 | GA [ga] | 788 |
| LU [ly] | 615 | SO [sɔ] | 492 | FU [fy] | 285 | BU [by] | 332 |

Table 2

*Descriptive statistics for all measures at T1, T2, and T3.*

| | Measures | *M* | *SD* | Skew[a] | Kurt[b] |
|---|---|---|---|---|---|
| T1 | Age (months) | 81.89 | 3.56 | 0.11 | -1.03 |
| | EVIP (/170) | 89.46 | 13.11 | -0.52 | 0.82 |
| | RCPM (/36) | 27.07 | 4.40 | -0.58 | -0.44 |
| | serial order STM (/81) | 45.10 | 8.05 | 0.25 | -0.47 |
| | filler (mean score) | .26 | .12 | 0.40 | -0.09 |
| | Hebb (mean score) | .38 | .24 | 0.63 | -0.67 |
| | filler (1st half) | .28 | .15 | - | - |
| | filler (2nd half) | .23 | .14 | - | - |
| | Hebb (1st half) | .33 | .23 | - | - |
| | Hebb (2nd half) | .43 | .27 | - | - |
| T2 | Age (months) | 93.71 | 3.58 | 0.13 | -1.02 |
| | EVIP (/170) | 96.40 | 18.48 | 0.00 | -0.70 |
| | RCPM (/36) | 28.96 | 4.16 | -0.40 | -0.68 |
| | serial order STM (/81) | 49.99 | 7.89 | 0.09 | -0.42 |
| | filler (mean score) | .28 | .15 | 0.84 | 0.62 |
| | Hebb (mean score) | .49 | .25 | 0.33 | -0.64 |
| | filler (1st half) | .31 | .19 | - | - |
| | filler (2nd half) | .26 | .16 | - | - |
| | Hebb (1st half) | .44 | .24 | - | - |
| | Hebb (2nd half) | .54 | .28 | - | - |
| | irregular word reading (/30) | 14.14 | 6.10 | -0.22 | -0.60 |
| | regular word reading (/30) | 25.08 | 4.07 | -1.48 | 3.05 |
| | pseudoword reading (/30) | 21.93 | 4.05 | -0.79 | 0.71 |
| T3 | Age (months) | 105.9 | 3.52 | 0.13 | -1.01 |
| | serial order STM (/81) | 55.35 | 6.7 | 0.06 | -0.11 |
| | filler (mean score) | .43 | .17 | 0.54 | -0.03 |
| | Hebb (mean score) | .67 | .23 | -0.43 | -0.90 |
| | filler (1st half) | .48 | .19 | - | - |
| | filler (2nd half) | .39 | .21 | - | - |
| | Hebb (1st half) | .60 | .24 | - | - |
| | Hebb (2nd half) | .73 | .25 | - | - |
| | irregular word reading (/30) | 20.38 | 5.36 | -0.55 | -0.07 |
| | regular word reading (/30) | 27.96 | 2.25 | -2.29 | 9.19 |
| | pseudoword reading (/30) | 24.62 | 3.16 | -0.94 | 1.19 |

*Note*. Hebb scores reflect proportions of correct responses. EVIP = vocabulary knowledge measure; RCPM = Raven’s colored progressive matrices.

[a] Standard error Skewness cutoff = + 0.22.

[b] Standard error Kurtosis cutoff = + 0.44.

Table 3

*Split-half reliability correlations for filler and Hebb measures across T1, T2, and T3.*

| Time point | Type | *r* | $BF_{10}$ |
|---|---|---|---|
| T1 | filler $1^{st}$ + $2^{nd}$ half | .54 | >100 |
| | Hebb $1^{st}$ + $2^{nd}$ half | .84 | >100 |
| T2 | filler $1^{st}$ + $2^{nd}$ half | .54 | >100 |
| | Hebb $1^{st}$ + $2^{nd}$ half | .82 | >100 |
| T3 | filler $1^{st}$ + $2^{nd}$ half | .59 | >100 |
| | Hebb $1^{st}$ + $2^{nd}$ half | .75 | >100 |

Table 4

*Summary table of the results of Bayesian regression analyses depicting, for each reading measure, the model associated with the largest evidence when contrasting serial order STM with LTL, while controlling for contemporaneous serial order STM and LTL measures as well as receptive vocabulary (EVIP) and nonverbal intelligence (RCPM).*

| | | |
|---|---|---|
| | **T1 → T2** | |
| **Irregular words** | **Regular words** | **Pseudowords** |
| Serial order STM | Serial order STM | Serial order STM |
| | **T1 → T3** | |
| **Irregular words** | **Regular words** | **Pseudowords** |
| Serial order STM | Serial order STM | Serial order STM |
| | **T2 → T3** | |
| **Irregular words** | **Regular words** | **Pseudowords** |
| T2 EVIP + contemp. serial order STM | Contemp. serial order LTL | Contemp. serial order STM |

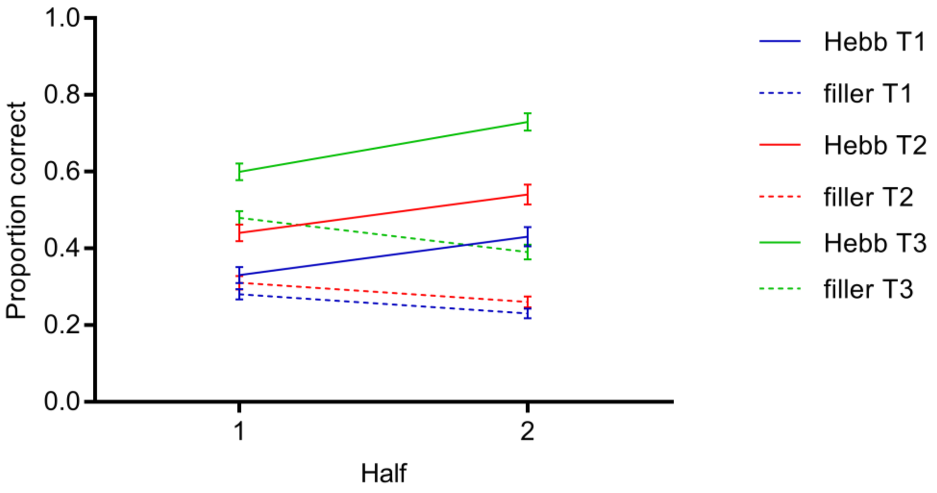


*Figure 1*. Mean proportion of items correctly recalled (with standard errors) at T1, T2, and T3 for Hebb and filler sequences by sequence halves.

# APPENDIX

Table A1

Correlations between all tasks for T1, T2, and T3.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 serial order STM_T1 | — | .37*** | .44*** | .35*** | .30* | .50*** | .38*** | .48*** | .34*** | .41*** | .45*** | .37*** | .36*** | .50*** | .43*** | .48*** | .24 | .43*** | .46*** | .39*** | .42*** |
| 2 Filler_Half2_T1 | | — | .31** | .20 | .24 | .32** | .22 | .29* | .20 | .18 | .31** | .27 | .11 | .27 | .32** | .34*** | .25 | .29* | .31** | .19 | .23 |
| 3 Hebb_Half2_T2 | | | — | .17 | .26 | .33** | .27 | .38*** | .15 | .20 | .21 | .16 | .06 | .33*** | .39*** | .33*** | .17 | .30** | .23 | .14 | .09 |
| 4 EVIP_T1 | | | | — | .23 | .26 | .35*** | .24 | .45*** | .31** | .25 | .15 | .09 | .32** | .15 | .15 | .45*** | .28* | .21 | .03 | .17 |
| 5 RCPM_T1 | | | | | — | .41*** | .17 | .11 | .25 | .53*** | .05 | .09 | .00 | .25 | .17 | .16 | .25 | .47*** | .12 | .02 | .08 |
| 6 serial order STM_T2 | | | | | | — | .32** | .28* | .16 | .30* | .29* | .21 | .18 | .45*** | .29* | .35*** | .13 | .35*** | .29* | .24 | .27 |
| 7 Filler_Half2_T2 | | | | | | | — | .30** | .24 | .20 | .15 | .23 | .17 | .44*** | .44*** | .26 | .19 | .26 | .15 | .05 | .22 |
| 8 Hebb_Half2_T2 | | | | | | | | — | .18 | .20 | .30* | .31** | .23 | .41*** | .32** | .43*** | .12 | .19 | .32** | .23 | .37*** |
| 9 EVIP_T2 | | | | | | | | | — | .36*** | .38*** | .26 | .05 | .27* | .17 | .33** | .57*** | .27* | .40*** | .21 | .21 |
| 10 RCPM_T2 | | | | | | | | | | — | .16 | .20 | .10 | .30* | .18 | .22 | .32** | .67*** | .21 | .13 | .23 |
| 11 Reading_IW_T2 | | | | | | | | | | | — | .73*** | .57*** | .36*** | .25 | .41*** | .40*** | .22 | .82*** | .65*** | .53*** |
| 12 Reading_RW_T2 | | | | | | | | | | | | — | .68*** | .33** | .28* | .37*** | .17 | .32** | .64*** | .60*** | .61*** |
| 13 Reading_PW_T2 | | | | | | | | | | | | | — | .34*** | .20 | .35*** | .07 | .25 | .51*** | .47*** | .55*** |
| 14 serial order STM_T3 | | | | | | | | | | | | | | — | .44*** | .33** | .27 | .22 | .43*** | .28* | .40*** |
| 15 Filler_Half2_T3 | | | | | | | | | | | | | | | — | .40*** | .17 | .26 | .26 | .20 | .12 |
| 16 Hebb_Half2_T3 | | | | | | | | | | | | | | | | — | .25 | .29* | .34*** | .32** | .25 |
| 17 EVIP_T3 | | | | | | | | | | | | | | | | | — | .22 | .39*** | .18 | .12 |
| 18 RCPM_T3 | | | | | | | | | | | | | | | | | | — | .18 | .11 | .25 |
| 19 Reading_IW_T3 | | | | | | | | | | | | | | | | | | | — | .69*** | .56*** |
| 20 Reading_RW_T3 | | | | | | | | | | | | | | | | | | | | — | .58*** |
| 21 Reading_PW_T3 | | | | | | | | | | | | | | | | | | | | | — |

* $BF_{10} > 10$, ** , $BF_{10} > 30$, *** $BF_{10} > 100$

IW = irregular words; RW = regular words; PW = pseudowords; RCPM = Raven's colored progressive matrices.

Table A2

Fixed-order regression analyses for **T1** serial order memory measures as predictors of T3 reading growth after controlling for contemporaneous serial order memory ability as well as nonverbal intelligence and vocabulary knowledge. $\Delta R^2$ are reported.

| | T3 irregular words | T3 regular words | T3 pseudowords |
|---|---|---|---|
| T2 irregular words | .686*** | | |
| T2 regular words | | .374*** | |
| T2 pseudowords | | | .273*** |
| T3 serial order STM | .021** | .004 | .059** |
| T3 2nd half Hebb | .001 | .008 | .001 |
| T1 RCPM | .001 | .007 | .001 |
| T1 EVIP | .002 | .008 | .003 |
| T1 serial order STM | .002 | .027** | .028* |
| T1 2nd half Hebb | .000 | .002 | .006 |

*** <.001, **<.01, *<.05.

Table A3

Fixed-order regression analyses for **T2** serial order memory measures as predictors of T3 reading growth after controlling for contemporaneous serial order memory ability as well as nonverbal intelligence and vocabulary knowledge. $\Delta R^2$ are reported.

| | T3 irregular words | T3 regular words | T3 pseudowords |
|---|---|---|---|
| T2 irregular words | .686*** | | |
| T2 regular words | | .374*** | |
| T2 pseudowords | | | .273*** |
| T3 serial order STM | .021** | .004 | .059** |
| T3 2nd half Hebb | .001 | .008 | .001 |
| T2 RCPM | .001 | .001 | .013 |
| T2 EVIP | .006 | .001 | .008 |
| T2 serial order STM | .000 | .005 | .006 |
| T2 2nd half Hebb | .001 | .001 | .025** |

*** <.001, **<.01, *<.05.